AD-A080 028 WHITE SANDS MISSILE RANGE NM F/G 12/1
APPLICATION OF ROBUST FILTERING AND SMOOTHING TO TRACKING DATA.(U)
DEC 79 W S AGEE, R H TURNER, J E GOMEZ
UNCLASSIFIED WSMR-TR-71 NL

1 OF 1
AD
A080028

END
DATE
FILMED
2-80
DDC

TECHNICAL REPORT

NO. 71

DECEMBER 1979

APPLICATION OF ROBUST FILTERING AND SMOOTHING TO TRACKING DATA

MATHEMATICAL SERVICES BRANCH
DATA SCIENCES DIVISION
US ARMY WHITE SANDS MISSILE RANGE
WHITE SANDS MISSILE RANGE, NEW MEXICO

TECHNICAL REPORT

NO. 71

Prepared by WILLIAM S. AGEE
Mathematician, Math Svcs Br

ROBERT H. TURNER
Mathematician, Math Svcs Br

JOSE E. GOMEZ
Mathematician, Math Svcs Br

Reviewed by JON E. GIBSON
Chief, Math Svcs Br

Approved by for PATRICK J. HIGGINS
Chief, Data Sciences Div

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)

Data Sciences Division
National Range Operations Directorate
White Sands Missile Range, New Mexico 88002

2a. REPORT SECURITY CLASSIFICATION

UNCLASSIFIED

2b. GROUP

NA

3. REPORT TITLE

(6) APPLICATION OF ROBUST FILTERING AND SMOOTHING TO TRACKING DATA

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

5. AUTHOR(S) (First name, middle initial, last name)

(10) William S. Agee, Robert H. Turner ~~and~~ Jose E. Gomez

6. REPORT DATE

(11) DECEMBER 1979

7a. TOTAL NO. OF PAGES

30

7b. NO. OF REFS

10

8a. CONTRACT OR GRANT NO.

b. PROJECT NO. (12) 34

c.

d.

9a. ORIGINATOR'S REPORT NUMBER(S)

(14) WSMR-TR-

(9) Technical Report No. 71

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10. DISTRIBUTION STATEMENT

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED.

11. SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY

13. ABSTRACT

Robust methods provide a fresh approach to the treatment of outliers or wild observations in filtering and smoothing applications. The robust M-estimates of regression are extended to filtering and fixed lag smoothing by employing a pseudo density of the observations in the derivations of the filter and fixed lag smoother. These robust methods are applied to tracking data to obtain improved estimation performance in the presence of wild observations. The improvement in estimation performance is evaluated via Monte Carlo using simulated tracking data.

DD FORM 1 NOV 65 1473 REPLACES DD FORM 1473, 1 JAN 64, WHICH IS OBSOLETE FOR ARMY USE.

344 800

UNCLASSIFIED

Security Classification

UNCLASSIFIED
Security Classification

<table>
<tr><th rowspan="2">14. KEY WORDS</th><th colspan="2">LINK A</th><th colspan="2">LINK B</th><th colspan="2">LINK C</th></tr>
<tr><th>ROLE</th><th>WT</th><th>ROLE</th><th>WT</th><th>ROLE</th><th>WT</th></tr>
<tr><td>Robust</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>M-estimates</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Filtering</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Smoothing</td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>

UNCLASSIFIED
Security Classification

ABSTRACT

Robust methods provide a fresh approach to the treatment of outliers or wild observations in filtering and smoothing applications. The robust M-estimates of regression are extended to filtering and fixed lag smoothing by employing a pseudo density of the observations in the derivations of the filter and fixed lag smoother. These robust methods are applied to tracking data to obtain improved estimation performance in the presence of wild observations. The improvement in estimation performance is evaluated via Monte Carlo using simulated tracking data.

TABLE OF CONTENTS

I. INTRODUCTION

Robust filtering and smoothing are a natural extension of the robust M-estimates of regression which have been developed by Huber [1]. The M-estimates of regression have been designed to perform well when the observations are contaminated by outliers. The conventional estimation techniques of least squares, minimum variance, maximum likelihood, etc. may become useless when the observations are contaminated by outliers. The robust M-estimates have been extremely successful in dealing with outliers in other data reduction problems, [2]. Outliers or wild data also present a problem when they occur in the observation sequence of a filter or smoother. These outliers have often been treated by testing the filter or smoother residuals. If the residual is too large as compared with some measure of dispersion of the residuals, the corresponding observation was rejected as being an outlier. Otherwise, the observation is processed normally by the filter or smoother. This procedure was often successful if only a small number of outliers were present but may breakdown when a larger proportion of outliers were present in the observation sequence. Also, in order for such an outlier detection method to be successful, a robust measure of dispersion of the filter or smoother residuals was necessary. These old methods of treating outliers in filters or smoother observations were added to the filter or smoother process as an afterthought. In contrast to this the development of robust filtering and smoothing methods by the use of M-estimates provides a method of treating outlying observations which is inherent in the filter or smoother equations.

Very little development has appeared on the application of robust estimation to filtering and smoothing. The most significant effort known to the author is the paper of Masreliez and Martin, [3]. Their development on the application of M-estimates to the Kalman filter is mainly theoretical. The emphasis in this report will be on the development of some practical results on the application of M-estimates to robust filtering and smoothing and the evaluation of these techniques for real and simulated tracking data.

The M-estimates for regression are discussed extensively in [1], [2], [4], and [5]. The following description provides a brief introduction to these robust regression estimates. Given scalar observations y_i, i=1,n of the linear model,

$$y_i = X_i\theta + e_i, \tag{1}$$

where X_i is a row vector of known independent variables and e_i is a random error term. We want to estimate the p-vector, θ. The M-estimate of θ minimizes

$$\sum_{i=1}^{n} \rho\left((y_i - X_i\theta)/s\right) , \tag{2}$$

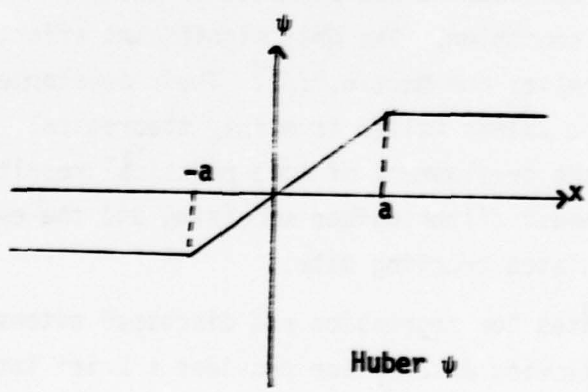
where $\rho(\cdot)$ is a specified function and s is a robust measure of the dispersion of the residuals, $y_i - X_i\theta$. Minimizing (2) by differentiating with respect to θ leads to

$$\sum_{i=1}^{n} X_i^T \psi\left((y_i - X_i\hat{\theta})/s\right) = 0 \tag{3}$$

where $\psi(\cdot)$ is the derivative of $\rho(\cdot)$ and $\hat{\theta}$ is the M-estimate of θ. (3) is the analog of the normal equations in least squares regression.

Rather than specifying the function ρ, M-estimates are usually described by specifying the function ψ. Several ψ functions have been proposed in the literature. These ψ functions may be grouped into two classes, the redescending type and the non-redescending type. The original ψ function proposed by Huber is of the non-redescending type and is given by

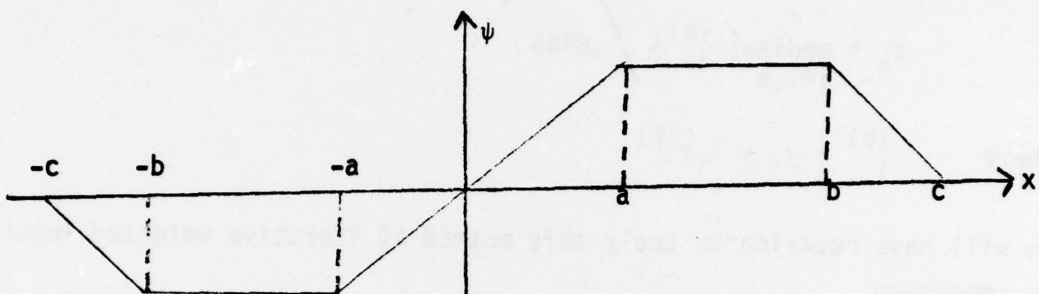
$$\psi(x) = \begin{cases} x & |x| \leq a \\ a \cdot \mathrm{sgn}(x) & |x| > a \end{cases} \tag{4}$$



Huber ψ

An example of the redescending ψ is furnished by the ψ function proposed by Hampel [6]. The Hampel ψ is given by

$$\psi(x) = \begin{cases} x & |x| \le a \\ a \operatorname{sgn}(x) & a < |x| \le b \\ a \left(\dfrac{x - c \operatorname{sgn}(x)}{b - c} \right) & b < |x| \le c \\ 0 & |x| > c \end{cases} \tag{5}$$



Hampel ψ

The Hampel ψ with breakpoints a,b,c, which we sometimes denote by Ha(a,b,c), is the *only ψ function* we will consider in this report. It is a very versatile function which can be made to take several shapes depending on the choice of the breakpoings. Besides using distinct breakpoints a,b,c, we also find the collapsed Hampel ψ's, Ha(a,a,a) and Ha(a,b,b) to be useful in filtering.

Since (3) is nonlinear, $\hat{\Theta}$ must be computed iteratively. A simple but highly effective method has been developed for the iterative solution of (3). This method is merely an iterative application of a weighted least squares algorithm, [2], [5], and [7]. Starting at an arbitrary point in the iteration sequence $\hat{\Theta}^{\{k\}}$, $\hat{\Theta}^{\{k+1\}}$ is computed by

$$\hat{\Theta}^{\{k+1\}} = \left(\sum_{j=1}^{n} W_j^{\{k\}} \; X_j^T X_j \right)^{-1} \sum_{i=1}^{n} W_i^{\{k\}} \; X_i^T y_i \tag{6}$$

(6) is easily recognized as the solution of a weighted least squares problem with weights $W_j^{\{k\}}$. The weights are computed iteratively by

$$W_j^{\{k\}} = \frac{\psi\left((y_j - X_j\hat{\Theta}^{\{k\}})/s_k\right)}{(y_j - X_j\hat{\Theta}^{\{k\}})/s_k} \tag{7}$$

A common dispersion measure s_k is

$$s_k = \underset{i=1,n}{\text{median}}\left|r_i^{\{k\}}\right| \Big/ .6745 \tag{8}$$

where $$r_i^{\{k\}} = y_i - X_i\hat{\Theta}^{\{k\}} \tag{9}$$

We will have occasion to apply this method of iterative weighted least squares to smoothing.

If the probability density function p of the measurement errors e_i is related to ψ by $p'/p=-\psi$, then the resulting M-estimate is maximum likelihood. If outliers are present in the observations, the probability density function should have heavier tails than the usual Gaussian density. For any ψ function, we call $p=e^{-\rho}$ a pseudo density. In the case of the Huber ψ, the pseudo density is an acceptable density having rather heavy tails. A plot of $e^{-\rho}$ for ψ=Hu(1), the Huber ψ function with breakpoint a=1, is given in Fig. 1. For a Hampel ψ function, Ha(a,b,c) the tails are so heavy that $e^{-\rho}$ is no longer a proper density function. A plot of $e^{-\rho}$ for ψ=Ha(1,2,3) is given in Fig. 2. Although the pseudo density may not be a proper probability density function, we derive some filters and smoothers in some of the conventional ways, e.g., conditional mean and maximizing the posterior density, with the probability densities of the measurement errors replaced by pseudo densities.

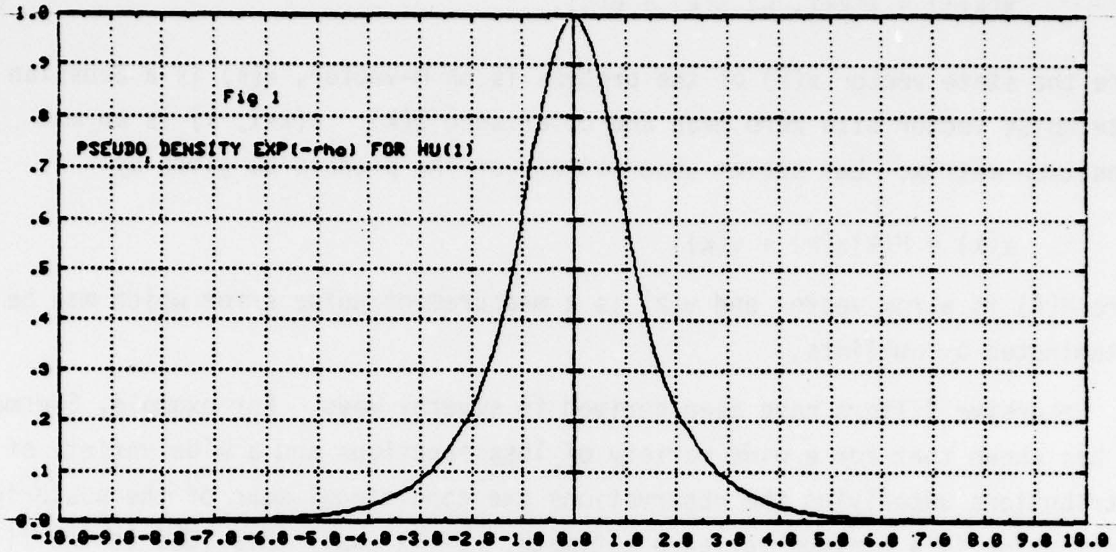


Fig 1

PSEUDO DENSITY EXP(-rho) FOR HU(1)

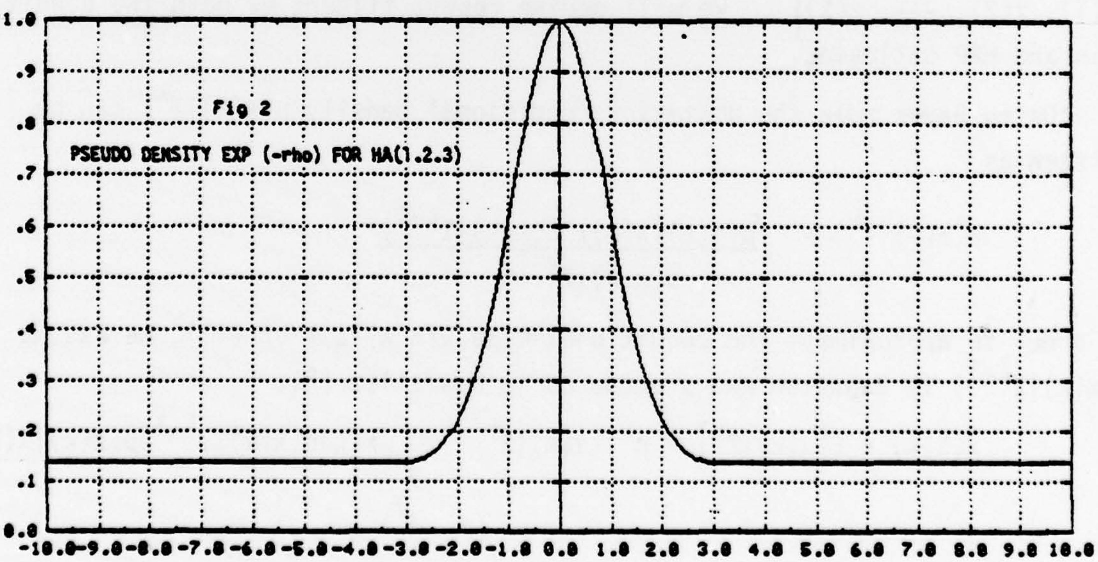


Fig 2

PSEUDO DENSITY EXP (-rho) FOR HA(1.2.3)

II. APPROXIMATE NON-GAUSSIAN FILTERS

Assume that the state of a process is given by the linear model

$$x(k+1) = \Phi(k+1, k)\, x(k) + u(k), \tag{10}$$

where the state vector $x(k)$ of the process is an m-vector, $u(k)$ is a Gaussian state noise vector with zero mean and covariance $Q(k)$. $\Phi(k+1, k)$ is an mxm transition matrix. Let scalar observations of the process be given by

$$z(k) = H(k)x(k) + v(k),$$

where $H(k)$ is a row vector and $v(k)$ is a measurement noise error which may be contaminated by outliers.

Recursive filters have been derived in several ways. For example, Sherman [8] has shown that for a wide variety of loss functions and a wide variety of distributions underlying the observations the conditional mean of the posterior distribution is a minimum variance estimate. If we apply this idea to the filtering case, we arrive at a minimum variance filter. Another method of recursive filter derivation is by maximizing the posterior distribution from which we obtain the MAP estimates. In either case we will work with the conditional probability density function $p(x(k)|Z^k)$, where Z^k denotes the set of observations, $Z^k = \left\{z(1), z(2), ---, z(k)\right\}$. We will derive robust filters by both the conditional mean and MAP estimates.

Using Bayes rule the posterior conditional density $p(x(k)|Z^k)$ can be written as

$$p(x(k)|Z^k) = \frac{p(z(k)|x(k))\; p(x(k)|Z^{k-1})}{p(z(k)|Z^{k-1})} \tag{11}$$

In order to approximate the conditional mean $E[x(k)|Z^k]$ of (11), we assume that $p(x(k)|Z^{k-1})$ is Gaussian and procede as in Masreliez [9].

$$x(k|k) = E\left[x(k)|Z^k\right] = p^{-1}(z(k)|Z^{k-1}) \int_{R^m} x(k)p(x(k)|Z^{k-1})p(z(k)|x(k))dx \tag{12}$$

Let $\hat{x}(k|k-1)$ be the mean of $p(x(k)|Z^{k-1})$ and let $P(k|k-1)$ be its covariance. Adding and subtracting $\hat{x}(k|k-1)$ to (12) gives

$$\hat{x}(k|k) = \hat{x}(k|k-1) + p^{-1}(z(k)|Z^{k-1}) \int_{R^m} (x(k) - \hat{x}(k|k-1))\, p(x(k)|Z^{k-1})p(z(k)|x(k))\, dx \tag{13}$$

Assuming $p(x(k)|Z^{k-1})$ to be Gaussian, we can write

$$(x(k) - \hat{x}(k|k-1)p(x(k)|Z^{k-1}) = -P(k|k-1) \frac{\partial}{\partial x(k)} p(x(k)|Z^{k-1}) \tag{14}$$

Then we can rewrite (13) as

$$\hat{x}(k|k) = \hat{x}(k|k-1) - p^{-1}(z(k)|Z^{k-1})P(k|k-1) \int_{R^m} \left[\frac{\partial}{\partial x(k)} p(x(k)|Z^{k-1})\right] p(z(k)|x(k))\, dx \tag{15}$$

Integrating (15) by parts,

$$\hat{x}(k|k) = \hat{x}(k|k-1) + p^{-1}(z(k)|Z^{k-1})P(k|k-1) \int_{R^m} p(x(k)|Z^{k-1}) \frac{\partial}{\partial x(k)} p(z(k)|x(k))\, dx \tag{16}$$

Using $\frac{\partial}{\partial x(k)} p(z(k)|x(k)) = -H^T(k) \frac{\partial}{\partial z(k)} p(z(k)|x(k))$ (17)

in (16) we obtain

$$\hat{x}(k|k) = \hat{x}(k|k-1) - p^{-1}(z(k)|Z^{k-1})P(k|k-1)H^T(k) \frac{\partial}{\partial z(k)} \int_{R^m} p(x(k)|Z^{k-1})p(z(k)|x(k))\, dx \tag{18}$$

Using $p(x(k)|Z^{k-1})p(z(k)|x(k)) = p(x(k),z(k)|Z^{k-1})$, (18) becomes

$$\hat{x}(k|k) = \hat{x}(k|k-1) + P(k|k-1)H^T(k)g(z(k)) \tag{19}$$

where $g(z(k))$ is the scalar

$$g(z(k)) = -p^{-1}(z(k)|Z^{k-1}) \frac{\partial}{\partial z(k)} p(z(k)|Z^{k-1}) \tag{20}$$

(19) and (20) were derived by Mareliez in [9] for approximating the minimum variance filter when measurement noise is non-Gaussian. In order to complete the specification of this approximate filter, it is necessary to derive an expression for the conditional second moment,

$$P(k|k) = E\left[(x(k) - \hat{x}(k|k)) (x(k) - \hat{x}(k|k))^T Z^k \right]$$

An expression for $P(k|k)$ is derived similar to the derivation for $\hat{x}(k|k)$. This derivation is outlined below.

$$\begin{aligned} P(k|k) = {} & E\left[(x(k) - \hat{x}(k|k-1)) (x(k) - \hat{x}(k|k-1))^T |Z^k \right] \\ & - (\hat{x}(k|k) - \hat{x}(k|k-1)) (\hat{x}(k) - \hat{x}(k|k-1))^T \end{aligned} \tag{21}$$

Let $S(k) = E\left[(x(k) - \hat{x}(k|k-1)) (x(k) - \hat{x}(k|k-1))^T |Z^k \right]$. Then using (11)

$$\begin{aligned} S(k) = {} & p^{-1}(z(k)|Z^{k-1}) \int_{R^m} (x(k) - \hat{x}(k|k-1)) (x(k) - \hat{x}(k|k-1))^T p(x(k)|Z^{k-1}) \\ & p(z(k)|x(k))\, dx \end{aligned} \tag{22}$$

Assuming $p(x(k)|Z^{k-1})$ Gaussian, we use (14) and integrate by parts twice to obtain,

$$S(k) = P(k|k-1) + P(k|k-1)H^T(k) \left\{ p^{-1}(z(k)|Z^{k-1}) \frac{\partial^2 p(z(k)|Z^{k-1})}{\partial z(k)^2} \right\} H(k)P(k|k-1) \tag{23}$$

Combining (23) with (21) and (19) gives

$$P(k|k) = P(k|k-1) - P(k|k-1)H^T(k)G(z(k))\, H(k)P(k|k-1) \tag{24}$$

where

$$G(z(k)) = \frac{\partial g(z(k))}{\partial z(k)} \tag{25}$$

A second method for approximate non-Gaussian filtering is the marginal maximum likelihood filter. This filter has also been called the maximum a posteriori or MAP filter. In this case we find the estimate $\hat{x}(k|k)$ which maximizes (11). Maximizing (11) is equivalent to minimizing $-\text{Log } p(x(k)|Z^k)$. Thus, the MAP estimate minimizes

$$L(x(k)) = -\log p(x(k)|Z^{k-1}) - \log p(z(k)|x(k)) \tag{26}$$

If we assume $p(x(k)|Z^{k-1})$ is Gaussian with mean $\hat{x}(k|k-1)$ and covariance $P(k|k-1)$, (26) becomes

$$L(x(k)) = \frac{1}{2} \; (x(k) - \hat{x}(k|k-1))^T \, P^{-1}(k|k-1) \; (x(k) - \hat{x}(k|k-1)) - \log p(z(k)|x(k)) \tag{27}$$

Minimizing (27) by setting $\frac{\partial L(x(k))}{\partial x(k)} = 0$ gives

$$\hat{x}(k|k) = \hat{x}(k|k-1) - P(k|k-1)H^T(k) \; \frac{\partial \log p(z(k)|\hat{x}(k|k))}{\partial z(k)} \tag{28}$$

The MAP filter formulation does not provide estimates of any of the moments of the density $p(x(k)|Z^k)$ but only the mode of the density. Thus, in order to continue the MAP filter from point to point, propagation of the first two moments of $p(x(k)|Z^k)$ in time must be provided by a different formulation. In other words the computation of the MAP estimate $\hat{x}(k+1|k+1)$ requires knowledge of the moments $\hat{x}(k+1|k)$ and $P(k+1|k)$ of $p(x(k+1)|Z^k)$ which are not available from the MAP formulation. We compute values of the moments $\hat{x}(k+1|k)$ and $P(k+1|k)$ from the conditional mean formulation given previously and consider the MAP estimate of (28) to be a correction to the conditional mean.

III. ROBUST CONDITIONAL MEAN FILTER

We apply (19), (20), and (24) to derive a robust filter based on the M-estimates. Given a ψ function for an M-estimate, we replace the density $p(z(k)|Z^{k-1})$ in (20) by the pseudo density to obtain a robust filter. Thus we use $p(z(k)|Z^{k-1}) = e^{-\rho}$ where ψ is the derivative of ρ. With this substitution the filter estimate of (19) becomes

$$\hat{x}(k|k) = \hat{x}(k|k-1) + \frac{P(k|k-1)H^T(k)}{s_k}\,\psi\left(\frac{z(k) - H(k)\hat{x}(k|k-1)}{s_k}\right) \tag{29}$$

The equation for the second moment given by (24) becomes

$$P(k|k) = P(k|k-1) - \left(\psi'\left(\frac{r(k)}{s_k}\right) \Big/ s_k^2\right) P(k|k-1)H^T(k)H(k)P(k|k-1), \tag{30}$$

where ψ' is the derivative of ψ and $r(k)=z(k)-H(k)\hat{x}(k|k-1)$ is the predicted residual. The filter equations are completed by the equations for the predicted moments,

$$\hat{x}(k+1|k) = \Phi(k+1,k)\hat{x}(k|k) \tag{31}$$

$$P(k+1|k) = \Phi(k+1,k)P(k|k)\Phi^T(k+1,k) + Q(k) \tag{32}$$

In order to insure the robustness of the filter of (29)-(32), the dispersion s_k of the predicted residuals must be specified so that it is insensitive to outliers. We use the MAD estimate of s_k computed from past residuals as

$$s_k = \underset{j=0,N-1}{\text{median}} \left|z(k-j)-H^T(k-j)\hat{x}(k-j|k-j-1)\right| \Big/ .6745 \tag{33}$$

where N is a suitably chosen integer. The robust conditional mean filter does not require iteration as was required for the M-estimates of regression. In this sense the conditional mean filter corresponds to the one step M-estimates described by Bickel [9].

IV. ROBUST MAP FILTER

As discussed previously the Non-Gaussian MAP filter merely provides a correction to the conditional mean filter. In the following we let $\hat{x}(k|k)$ denote the MAP estimate and $\bar{x}(k|k)$ the approximate mean of $p(x(k)|Z^k)$. Given a ψ function for an M-estimate, we replace the density $p(z(k)|\hat{x}(k|k))$ in (28) by the pseudo density to obtain the robust MAP estimate. Thus, we use $p(z(k)|\hat{x}(k|k)) = e^{-\rho}$ where ψ is the derivative of ρ. With this substitution the MAP estimate of (28) becomes

$$\hat{x}(k|k) = \bar{x}(k|k-1) + \frac{P(k|k-1)H^T(k)}{s_k}\ \psi\left(\frac{z(k) - H(k)\hat{x}(k|k)}{s_k}\right) \tag{34}$$

To complete the description of this robust filter we use the conditional moments

$$\bar{x}(k+1|k) = \Phi(k+1,k)\bar{x}(k|k) \tag{35}$$

$$\bar{x}(k|k) = \bar{x}(k|k-1) + \frac{P(k|k-1)H^T(k)}{s_k}\ \psi\left(\frac{z(k) - H(k)\bar{x}(k|k-1)}{s_k}\right) \tag{36}$$

and

$$P(k|k) = P(k|k-1) - \left(\psi'\ \frac{r(k)}{s_k} \Big/ s_k^2\right) P(k|k-1)H^T(k)H(k)P(k|k-1) \tag{37}$$

where $r(k) = z(k) - H(k)\bar{x}(k|k-1)$

also $P(k+1|k) = \Phi(k+1,k)P(k|k)\Phi^T(k+1,k) + Q(k)$ (38)

Again we use the MAD estimate on past residuals

$$s_k = \underset{j=0,N-1}{\text{median}} \left|z(k-j)-H(k-j)\bar{x}(k-j|k-j)\right| \Big/ .6745 \tag{39}$$

Note that the robust MAP estimate specified by (34) requires iteration for solution since $\hat{x}(k|k)$ appears nonlinearly on the right hand side of (34). Several simple methods for iteration are readily apparent. The first method is to iterate

(34) directly. Let $\hat{x}^{(\alpha)}(k|k)$ be an arbitrary point in the iteration sequence and let $\hat{x}^{(0)}(k|k) = \overline{x}(k|k-1)$. Then we iterate (34) as

$$\hat{x}^{(\alpha+1)}(k|k) = \overline{x}(k|k-1) + \frac{P(k|k-1)H^T(k)}{s_k}\,\psi\left(\frac{z(k) - H(k)\hat{x}^{(\alpha)}(k|k)}{s_k}\right) \qquad (40)$$

In this case we can replace (35) by

$$\overline{x}(k+1|k) = \Phi(k+1,k)\hat{x}^{(1)}(k|k) \qquad (41)$$

Another simple method for iteratively solving (34) is similar to the iterative weighted least squares method used to solve the robust regression equations. If $\hat{x}^{(\alpha)}(k|k)$ is an arbitrary point in the iteration sequence, this method solves (34) iteratively as

$$\hat{x}^{(\alpha+1)}(k|k) = \overline{x}(k|k-1) + W_k^{(\alpha)}\,\frac{P^{(\alpha)}(k)\,H^T(k)}{s_k^2}\,(z(k) - H(k)\overline{x}(k|k-1)), \qquad (42)$$

where $W_k^{(\alpha)}$ is a scalar weight, $0 \le W_k^{(\alpha)} \le 1$ and $\hat{x}^{(0)}(k|k) = \overline{x}(k|k-1)$

$$W_k^{(\alpha)} = \frac{\psi\left(\dfrac{z(k) - H(k)\hat{x}^{(\alpha)}(k|k)}{s_k}\right)}{\dfrac{z(k) - H(k)\hat{x}^{(\alpha)}(k|k)}{s_k}} \qquad (43)$$

and

$$P^{(\alpha)}(k) = \left[P^{-1}(k|k-1) + \frac{W_k^{(\alpha)}}{s_k^2}\,H^T(k)H(k)\right]^{-1} \qquad (44)$$

The filter correction provided by (34) to the conditional mean filter can be quite sizeable. A change to the prediction equation which uses $\hat{x}(k|k)$ on the right hand of (41) rather than $\hat{x}^{(1)}(k|k)$ might be a desirable in some applications.

V. APPROXIMATE NON-GAUSSIAN SMOOTHING

In the following some robust fixed lag smoothing equations are derived in a similar manner to the derivation of the robust filter equations, i.e., using the conditional mean and MAP formulations. In fixed lag smoothing an estimate of the state $x(k)$ of the system described by (10) and (11) is desired using the measurements $z(1), z(2), ---, z(k), z(k+1), ---, z(k+N)$. Thus, there is a lag of N points in obtaining this smoothed state estimate. Let $\Delta Z^{k+N} = \left\{ z(k+1), z(k+2), ---, z(k+N) \right\}$. Then $Z^{k+N} = Z^k \cup \Delta Z^{k+N}$. The posterior conditional desnity $p(x(k)|Z^{k+N})$ is given by

$$p(x(k)|Z^{k+N}) = \frac{p(\Delta Z^{k+N}|x(k))\ p(x(k)|Z^k)}{p(\Delta Z^{k+N}|Z^k)} \tag{45}$$

We assume that the density $p(x(k)|Z^k)$ is Gaussian and procede as before to obtain the conditional mean, $\hat{x}(k|k+N) = E\left[x(k)|Z^{k+N} \right]$.

The conditional mean is given by

$$\hat{x}(k|k+N) = p^{-1}(\Delta Z^{k+N}|Z^k) \int_{R^m} x(k)\ p(\Delta Z^{k+N}|x(k))\ p(x(k)|Z^k)\ dx(k) \tag{46}$$

Adding and subtracting $\hat{x}(k|k)$ to (46) and assuming that $p(x(k)|Z^k)$ is Gaussian gives

$$\hat{x}(k|k+N) = \hat{x}(k|k) - p^{-1}(\Delta Z^{k+N}|Z^k)\ P(k|k) \int_{R^m} \left[\frac{\partial}{\partial x(k)}\ p(x(k)|Z^k)\right] p(\Delta Z^{k+N}|x(k))\ dx(k) \tag{47}$$

where $P(k|k)$ is the covariance of $p(x(k)|Z^k)$. Integrating by parts

$$\hat{x}(k|k+N) = \hat{x}(k|k) + p^{-1}(\Delta Z^{k+N}|Z^k)\ P(k|k) \int_{R^m} \left[\frac{\partial}{\partial x(k)}\ p(\Delta Z^{k+N}|x(k))\right] p(x(k)|Z^k)\ dx(k) \tag{48}$$

Assuming statistical independence of the observations

$$p(\Delta Z^{k+N}|x(k)) = \prod_{j=1}^{N} p(z(k+j)|x(k)) \tag{49}$$

Then

$$\frac{\partial}{\partial x(k)} p(\Delta Z^{k+N}|x(k)) = \sum_{j=1}^{N} Q_j \frac{\partial p(z(k+j)|x(k))}{\partial x(k)} \tag{50}$$

where

$$Q_j = \prod_{i \neq j} p(z(k+i)|x(k)) \tag{51}$$

Also,

$$\frac{\partial p(z(k+j)|x(k))}{\partial x(k)} = -\Phi^T(t_{k+j}, t_k)\, H^T(k+j) \frac{\partial p(z(k+j)|x(k))}{\partial z(k+j)} \tag{52}$$

Substituting (49)-(52) in (48)

$$\hat{x}(k|k+N) = \hat{x}(k|k) + p^{-1}(\Delta Z^{k+N}|Z^k)\, P(k|k) \sum_{j=1}^{N} \Phi^T(t_{k+j}, t_k) \tag{53}$$

$$H^T(k+j) \frac{\partial}{\partial z(k+j)} \int_{R^m} p(\Delta Z^{k+N}|x(k)) p(x(k)|Z^k)\, dx(k)$$

But

$$p(\Delta Z^{k+N}|x(k)) p(x(k)|Z^k) = p(\Delta Z^{k+N}, x(k)|Z^k) \tag{54}$$

Substituting (54) into (53) yields

$$\hat{x}(k|k+N) = \hat{x}(k|k) - P(k|k) \sum_{j=1}^{N} \Phi^T(t_{k+j}, t_k)\, H^T(k+j)\, p^{-1}(z(k+j)|Z^k) \frac{\partial p(z(k+j)|Z^k)}{\partial z(k+j)} \tag{55}$$

The filtered estimate $\hat{x}(k|k)$ and $P(k|k)$ are given by (19) and (24).

A slightly different robust fixed lag smoother is obtained if, instead of finding the conditional mean of $p(x(k)|Z^{k+N})$, we compute the mode of $p(x(k)|Z^{k+N})$. We will call this method the marginal maximum likelihood or the MAP formulation of the robust smoother. This estimate maximizes (45) or equivalently minimizes

$$L(x(k)) = -\log p(\Delta Z^{k+N}|x(k)) - \log p(x(k)|Z^k) \tag{56}$$

Assuming that $p(x(k)|Z^k)$ is Gaussian with mean $\hat{x}(k|k)$ and covariance $P(k|k)$ (56) becomes,

$$L(x(k)) = 1/2\ (x(k) - \hat{x}(k|k))^T P^{-1}(k|k)\ (x(k) - \hat{x}(k|k)) - \log p(\Delta Z^{k+N}|x(k)) \tag{57}$$

Minimizing (57) by setting $\frac{\partial L(x(k))}{\partial x(k)} = 0$ gives

$$\hat{x}(k|k+N) = \hat{x}(k|k) + P(k|k)\ p^{-1}(\Delta Z^{k+N}|x(k))\ \frac{\partial p(\Delta Z^{k+N}|x(k))}{\partial x(k)} \tag{58}$$

Assuming statistical independence of the observations, we use (49) and (52) in (58) to obtain

$$\hat{x}(k|k+N) = \hat{x}(k|k) - P(k|k) \sum_{j=1}^{N} \Phi^T(t_{k+j}, t_k)\ H^T(k+j)\ p^{-1}(z(k+j)|\hat{x}(k|k+N)) \tag{59}$$

$$\frac{\partial p(z(k+j)|\hat{x}(k|k+N))}{\partial \hat{x}(k|k+N)}$$

$\hat{x}(k|k)$ and $P(k|k)$ are obtained via (19) and (24). The MAP formulation of (59) can be viewed as a correction to the conditional mean formulation of (55). Since $\hat{x}(k|k+N)$ appears nonlinearly on the right hand side of (59), the solution of (59) will require iteration.

VI. ROBUST CONDITIONAL MEAN SMOOTHING

In order to obtain a robust fixed lag smoother via the conditional mean formulation, we replace the densities, $p(z(k+j)|Z^k)$ in (55) by pseudo densities. Thus, given a ψ function for an M-estimate, we use $p(z(k+j)|Z^k) = e^{-\rho}$ where ψ is the derivative of ρ. With this substitution the robust smoothing equation of (55) becomes

$$\hat{x}(k|k+N) = \hat{x}(k|k) + P(k|k) \sum_{j=1}^{N} \frac{\Phi^T (t_{k+j}, t_k) H^T(k+j)}{s_{k+j}} \psi\left(\frac{z(k+j) - H(k+j)\,\Phi(t_{k+j}, t_k)\,\hat{x}(k|k)}{s_{k+j}}\right) \tag{60}$$

No iteration is required for the computation in (60) so that this estimate corresponds to the one step M-estimate.

In order to insure the robustness of the estimate in (60), it is necessary that s_{k+j} in (60) be a robust measure of dispersion of the residuals, $r(k+j) = z(k+j) - H\Phi(t_{k+j}, t_j)\,\hat{x}(k|k)$. Several alternatives are possible for computing s_{k+j}, depending on the assumptions made about the statistics of the residuals. The simplest method for computing s_{k+j} is to assume it is independent of j. In this case we can use the same estimate (39) as the filter. Another possible estimate when assuming s_{k+j} to be independent of j is

$$s_{k+j} = s_k = \operatorname*{median}_{j=1,N-1} |z(k+j) - H(k+j)\,\Phi(t_{k-1+j}, t_{k-1})\,\hat{x}(k-1|k-1+N)| \Big/ .6745 \tag{61}$$

If an estimate s_{k+j} is required for each j, then values of the forward residuals must be saved for past values of k. In this case we might use

$$s_{k+j} = \operatorname*{median}_{\ell=1,Q} |z(k-\ell+j) - H(k-\ell+j)\,\Phi(t_{k-\ell+j}, t_{k-\ell})\,\hat{x}(k-\ell|k-\ell+N)| \Big/ .6745 \tag{62}$$

where Q is a suitable integer.

VII. ROBUST MAP SMOOTHING

To obtain a Robust fixed lag smoother from the MAP smoothing formulation of (59), the densities $p(z(k+j)|Z^k)$ are replaced by pseudo densities, $p(z(k+j)|Z^k) = e^{-\rho}$, where the derivative of ρ is the influence function ψ for an M-estimate. With this substitution (59) is

$$\hat{x}(k|k+N) = \hat{x}(k|k) + P(k|k) \sum_{j=1}^{N} \frac{\Phi^T(t_{k+j}, t_k)\, H^T(k+j)}{s_{k+j}}$$

$$\psi\left(\frac{z(k+j) - H(k+j)\,\Phi(t_{k+j}, t_k)\,\hat{x}(k|k+N)}{s_{k+j}}\right) \tag{63}$$

Again s_{k+j} is a robust measure of dispersion of the residuals, $r(k+j) = z(k+j) - H(k+j)\,\Phi(t_{k+j}, t_k)\,\hat{x}(k|k+N)$. Since $\hat{x}(k|k+N)$ appears nonlinearly on the right hand side of (63), iteration is required for its solution. The most obvious method is to iterate (63) in its present form, i.e., to replace $\hat{x}(k|k+N)$ on the right hand side of (63) by $\hat{x}^{(\alpha)}(k|k+N)$ and $\hat{x}(k|k+N)$ on the left side of (63) by $\hat{x}^{(\alpha+1)}(k|k+N)$. This iteration is generally unstable and therefore unuseable. A much better iteration procedure for (63) is to use weights in a manner similar to the iterated weighted least squares of (6) and (7) used to obtain the M-estimates of regression. Using this procedure the iterative solution of (63) is defined by

$$\hat{x}^{(\alpha+1)}(k|k+N) = \hat{x}(k|k) + P^{(\alpha)}(k) \sum_{j=1}^{N} \frac{W_j^{(\alpha)}\, \Phi^T(t_{k+j}, t_k)\, H^T(k+j)}{s_{k+j}^2}$$

$$(z(k+j) - H(k+j)\,\Phi(t_{k+j}, t_k)\,\hat{x}(k|k)) \tag{64}$$

where

$$P^{(\alpha)}(k) = \left[P^{-1}(k|k) + \sum_{j=1}^{N} \frac{W_j^{(\alpha)}}{s_{k+j}^2}\, \Phi^T(t_{k+j}, t_k)\, H^T(k+j)H(k+j)\,\Phi(t_{k+j}, t_k)\right]^{-1} \tag{65}$$

and

$$W_j^{(\alpha)} = \frac{\psi\left(\dfrac{z(k+j) - H(k+j)\Phi\left(t_{k+j},\ t_k\right)\ \hat{x}^{(\alpha)}(k|k+N)}{s_{k+j}}\right)}{\dfrac{z(k+j) - H(k+j)\Phi\left(t_{k+j},\ t_k\right)\ \hat{x}^{(\alpha)}(k|k+N)}{s_{k+j}}} \tag{66}$$

The iteration starts with $\hat{x}^{(0)}(k|k+N) = \hat{x}(k|k)$. The filtered quantities $\hat{x}(k|k)$ and $P(k|k)$ are obtained from (29) and (30). The iteration specified by (64)-(66) usually converges very rapidly, most often in two or three cycles.

VIII. EVALUATION OF ROBUST FILTERING

Evaluation of the robust filtering methods described here has been done with a view toward eventual application to trajectory estimation. Emphasis in the evaluation is on the use of simulated rather than real trajectory data. This allows a quantitative determination of any advantages in the use of robust filtering in the presence of outliers and also any loss in efficiency using robust methods when no outliers are present. The simulated trajectory is that of a constant velocity, level flying aircraft. The filter model assumes the trajectory to have constant acceleration in three cartesian coordinates. Let x, y, z be the East, North, and Up components of trajectory position. We assume that the dynamic model for each of the coordinates is given by

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} 1 & \Delta & \Delta^2/2 \\ 0 & 1 & \Delta \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ w(k) \end{bmatrix} \tag{67}$$

where $\Delta = t_{k+1} - t_k$ so that $x_1(k)$, $x_2(k)$, and $x_3(k)$ are position, velocity, and acceleration components, respectively. $w(k)$ is a zero mean Gaussian acceleration state noise with variance q. The filter observations, $z(k)$ are scalar positions corrupted by additive noise.

$$z(k) = Hx(k) + v(k) \tag{68}$$

with $H = [1 \quad 0 \quad 0]$. The measurement noise $v(k)$ is Gaussian with covariance $R(k)$. Two different methods are used to generate outliers in the observations. One method is to choose the variance $R(k)$ of the noise $v(k)$ as

$$R(k) = \begin{cases} R_1(k) & \text{if no outlier} \\ R_2(k) & \text{if outlier present} \end{cases} \tag{69}$$

with $R_2(k) >> R_1(k)$. In this case the mean of the measurement noise is zero. The second method used to generate outliers in the observations is to choose the mean $\mu(k)$ of the observation noise as

$$\mu(k) = \begin{cases} 0 & \text{if no outlier} \\ \mu_1(k) & \text{if outlier present} \end{cases} \tag{70}$$

In this case the variance of the observation noise is $R_1(k)$. In either method we must decide whether or not an outlier is to be present in the data at each time t_k. We do this by using a two state Markov chain. Let i denote the state of the Markov chain. i = 1 is the state of no outlier present and i = 2 is the state if an outlier is present in the data. Let $P_{ij}(k)$ be the probability of a transition from state i to state j in the interval (t_{k-1}, t_k). The transition probabilities are chosen to provide a given percentage of outliers and desired run lengths of outliers in the observations. The transitions between states are realized by use of a pseudo random number generator.

The constant velocity trajectory used for evaluation is given by

$$\begin{aligned} x(t_{k+1}) &= x(t_k) + \dot{x}(t_{k+1} - t_k) \\ y(t_{k+1}) &= y(t_k) + \dot{y}(t_{k+1} - t_k) \\ z(t_{k+1}) &= z(t_k) + \dot{z}(t_{k+1} - t_k) \end{aligned} \tag{71}$$

with $\dot{x} = -550$ft/sec, $\dot{y} = -252$ft/sec, and $\dot{z} = 0$. A sampling interval of $t_{k+1} - t_k =$ .05 sec was used. a monte carlo evaluation of the robust filter is done by computing some statistics of the filtering errors over N filter runs. Let $\hat{x}_i(t_k)$, $\hat{y}_i(t_k)$, and $\hat{z}_i(t_k)$ denote the filtered position estimates at time t_k for the $i^{\underline{th}}$ run and let $\tilde{x}_i(t_k) = \hat{x}_i(t_k) - x(t_k)$, $\tilde{y}_i(t_k) = \hat{y}_i(t_k) - y(t_k)$, and $\tilde{z}_i(t_k) = \hat{z}_i(t_k) - z(t_k)$ denote the errors in the filtered positions for the $i^{\underline{th}}$ filter run at time t_k. Also, let $\tilde{\dot{x}}_i(t_k) = \hat{\dot{x}}_i(t_k) - \dot{x}$, $\tilde{\dot{y}}_i(t_k) = \hat{\dot{y}}_i(t_k) - \dot{y}$, and $\tilde{\dot{z}}_i(t_k) = \hat{\dot{z}}_i(t_k) - \dot{z}$ denote errors in filtered velocity estimates and $\tilde{\ddot{x}}_i(t_k) = \hat{\ddot{x}}_i(t_k) - \ddot{x}$, $\tilde{\ddot{y}}_i(t_k) = \hat{\ddot{y}}_i(t_k) - \ddot{y}$, and $\tilde{\ddot{z}}_i(t_k) = \hat{\ddot{z}}_i(t_k) - \ddot{z}$ be the errors in the filtered accelerations for the $i^{\underline{th}}$ run at time t_k. We evaluate only the RSS position, velocity, and acceleration estimates,

$$R_i(t_k) = \left(\tilde{x}_i^2(t_k) + \tilde{y}_i^2(t_k) + \tilde{Z}_i^2(t_k)\right)^{1/2}$$

$$\dot{R}_i(t_k) = \left(\dot{\tilde{x}}_i^2(t_k) + \dot{\tilde{y}}_i^2(t_k) + \dot{\tilde{Z}}_i^2(t_k)\right)^{1/2}$$

$$\ddot{R}_i(t_k) = \left(\ddot{\tilde{x}}_i^2(t_k) + \ddot{\tilde{y}}_i^2(t_k) + \ddot{\tilde{Z}}_i^2(t_k)\right)^{1/2} \quad (72)$$

We compute the sample averages of the RSS errors defined in (72).

$$\overline{R}(t_k) = \frac{1}{N}\sum_{i=1}^{N} R_i(t_k)$$

$$\overline{\dot{R}}(t_k) = \frac{1}{N}\sum_{i=1}^{N} \dot{R}_i(t_k)$$

$$\overline{\ddot{R}}(t_k) = \frac{1}{N}\sum_{i=1}^{N} \ddot{R}_i(t_k) \quad (73)$$

In order to reduce the evaluation of the robust filter to the comparison of only a few numbers, we compute the time average of estimation errors in (73)

$$\overline{\overline{R}} = \frac{1}{M}\sum_{k=1}^{M} \overline{R}(t_k)$$

$$\overline{\overline{\dot{R}}} = \frac{1}{M}\sum_{k=1}^{M} \overline{\dot{R}}(t_k)$$

$$\overline{\overline{\ddot{R}}} - \frac{1}{M}\sum_{k=1}^{M} \overline{\ddot{R}}(t_k) \quad (74)$$

where M is the total number of filtered points. Unless otherwise specified,

$N = 25$, $\sqrt{R_1(k)} = 20$ ft., $P_{12}(k) = .05$, and $P_{21}(k) = .5$. These transition probabilities for the Markov chain provide an outlier contamination of about 9% and an average outlier run length of three. A state noise covariance, $q = 5$, was used for all filter runs. In all runs of the robust filter the measurement noise covariance, $R_1(k)$, was unknown to the filter. The residual variance, which is the only quantity required by the filter which involves $R_1(k)$, was estimated using the MAD estimate of (33).

Figure 3 compares the average RSS position errors for the two robust bilters using the Hampel ψ functions $H_a(4, 4, 4)$ and $H_a(2, 3, 4)$. Figure 4 gives the RSS velocity error comparison for the same two filters. Also indicated in Figures 3 and 4 are the ideal RSS error values which were obtained using an ordinary Kalman filter with no outliers present and using a known measurement covariance, $R_1(k) = 400$. We note from Figures 3 and 4 that neither of the robust filters lose much efficiency from the ideal values when no outliers are present. In Figures 3 and 4 the magnitude of the outlier contamination is $\mu_1(k) = C \cdot \sqrt{R_1(k)}$ for various values of C. The error curves in Figures 3 and 4 behave as expected. Since outliers small in relation to the measurement noise are hardest to detect, the error curve rises sharply. Outliers large relative to the measurement noise are easy to detect so that the error curve returns to zero for large outliers. We see from Figures 3 and 4 that $H_a(2, 3, 4)$ has a significantly smaller estimation error than $H_a(4, 4, 4)$. Except for the way in which the dispersion of the residuals is measured, i.e., the MAD estimate of (33), $H_a(4, 4, 4)$ is a conventional way of handling outliers in a Kalman filter application. Using $H_a(4, 4, 4)$ any observation whose predicted residual is greater than $4 \cdot S_k$ is not processed and any observation whose predicted residual is less than $4 \cdot S_k$ is processed as an ordinary Kalman filter observation.

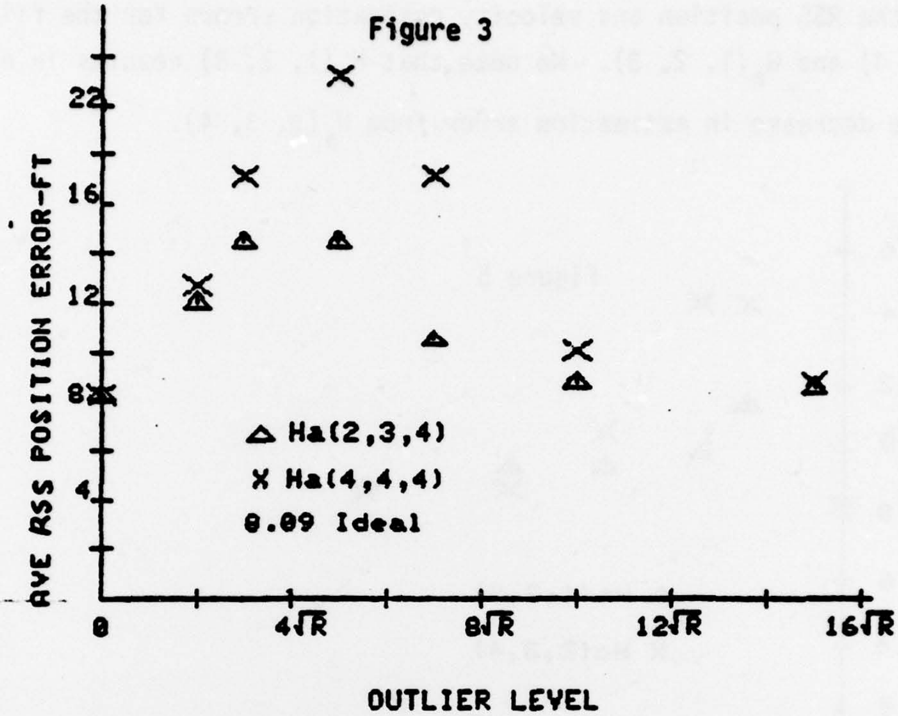


Figure 3
AVE RSS POSITION ERROR-FT
OUTLIER LEVEL
△ Ha(2,3,4)
X Ha(4,4,4)
8.09 Ideal
0
4√R
8√R
12√R
16√R
4
8
12
16
22

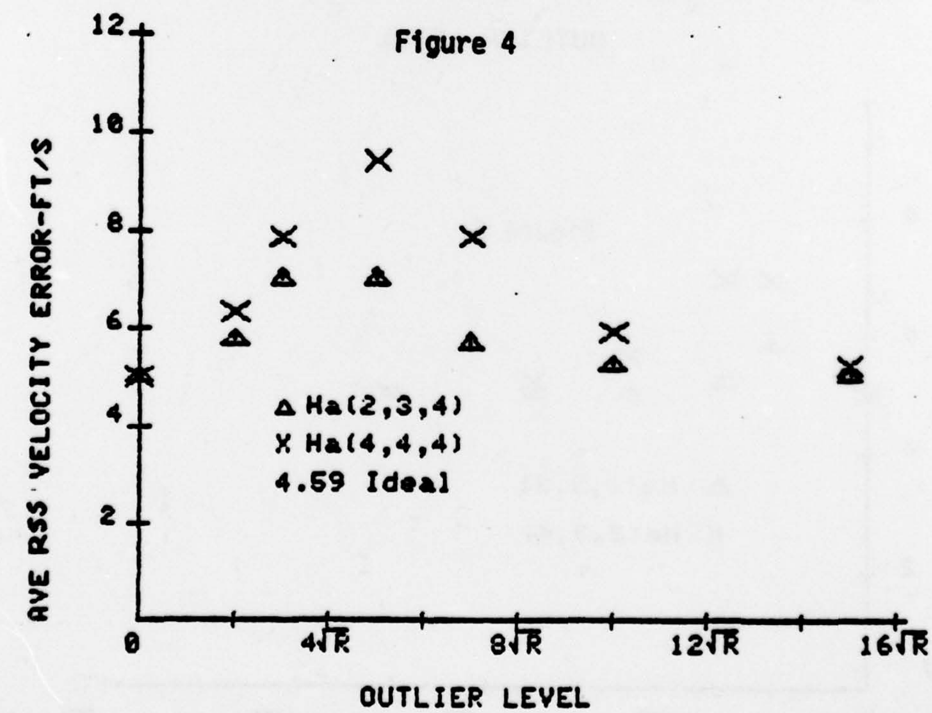
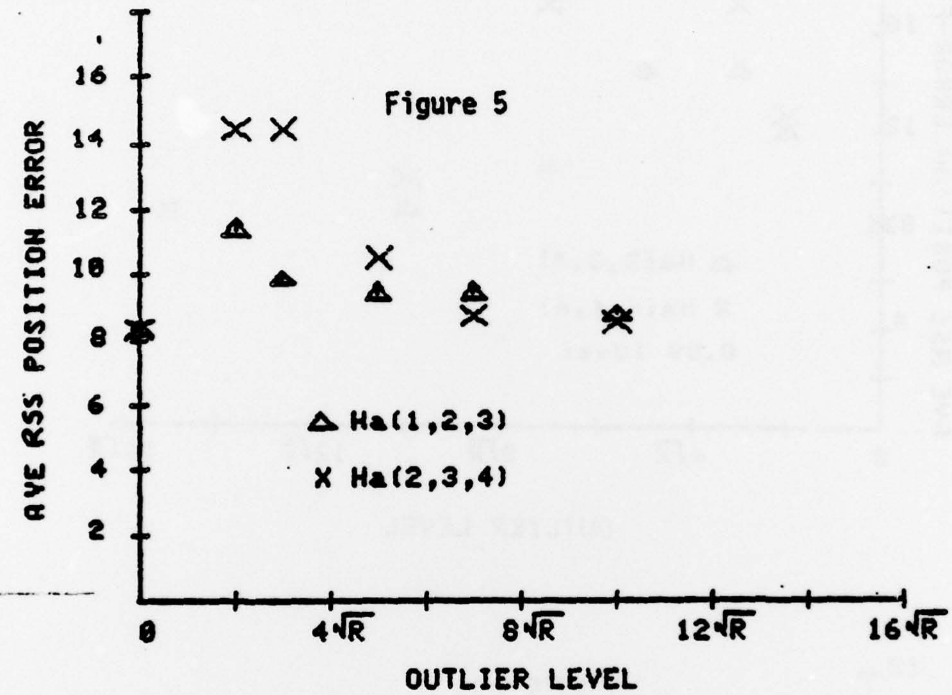
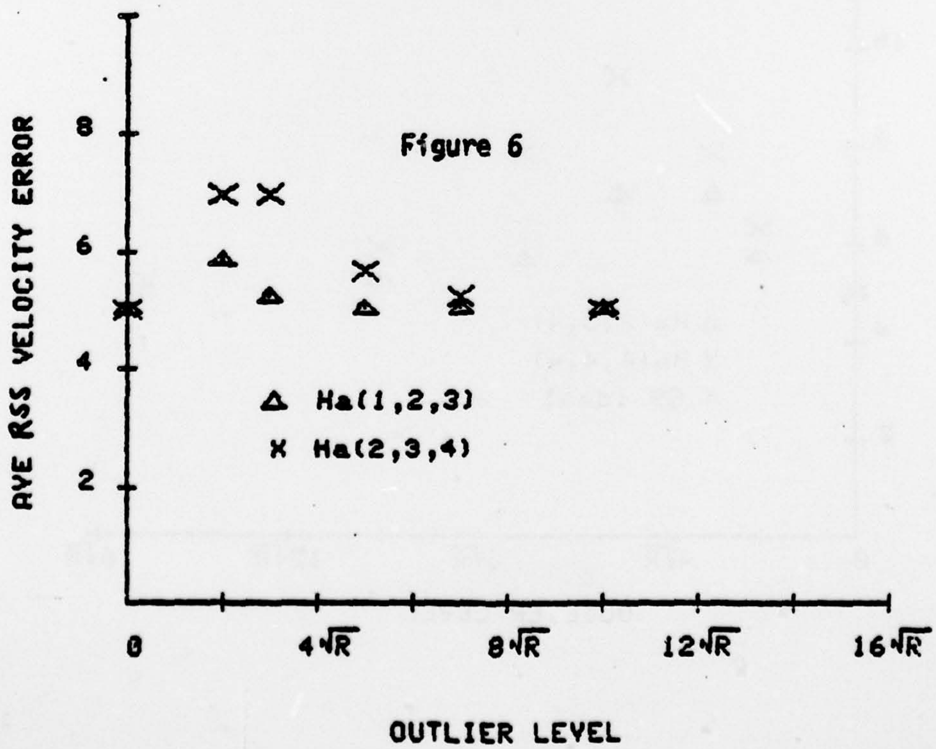


Figure 4
AVE RSS VELOCITY ERROR-FT/S
OUTLIER LEVEL
△ Ha(2,3,4)
X Ha(4,4,4)
4.59 Ideal
0
4√R
8√R
12√R
16√R
2
4
6
8
10
12

We can reduce the RSS estimation errors for the robust filter still farther by pulling in the breakpoints of the Hampel ψ function. Figures 5 and 6 compare the RSS position and velocity estimation errors for the filters $H_a(2, 3, 4)$ and $H_a(1, 2, 3)$. We note that $H_a(1, 2, 3)$ results in a considerable decrease in estimation error from $H_a(2, 3, 4)$.





The iterated filter or MAP filter given in (34) - (38) was evaluated under the same conditions as the conditional mean filter. Comparison of the average RSS estimation errors for the conditional mean and MAP filters showed no discernible differences. Because the MAP filter offers no improvement over the conditional mean filter and because the conditional mean filter is the simpler of the two robust filters, we will not give any further evaluation of the MAP filter.

The robust filter was also evaluated with respect to its ability to adapt to changes in measurement noise variance. Using the same simulated trajectory as before, the measurement noise standard deviation was $\sqrt{R_1(k)}$ = 20 ft for $0 \leq t \leq 25$ sec., $\sqrt{R_1(k)}$ = 100 ft for $25 < t \leq 40$ sec., and $\sqrt{R_1(k)}$ = 50 ft for $40 < t \leq 50$ sec. Figures 7 and 8 compare the RSS position and velocity estimation errors for the filters using $H_a(1, 2, 3)$ and $H_a(3, 3, 3)$ for the above measurement noise variations. Again, we find that the robust filter $H_a(1, 2, 3)$ has a considerably smaller mean square error in the presence of outliers than the more conventional way of handling outliers implemented in $H_a(3, 3, 3)$.

Figure 7

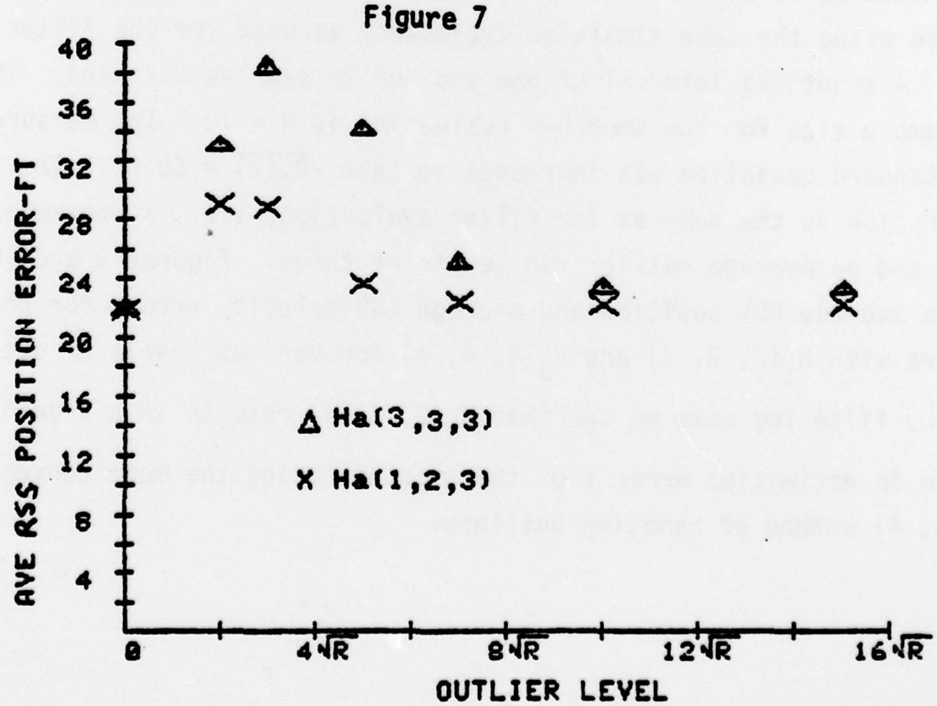
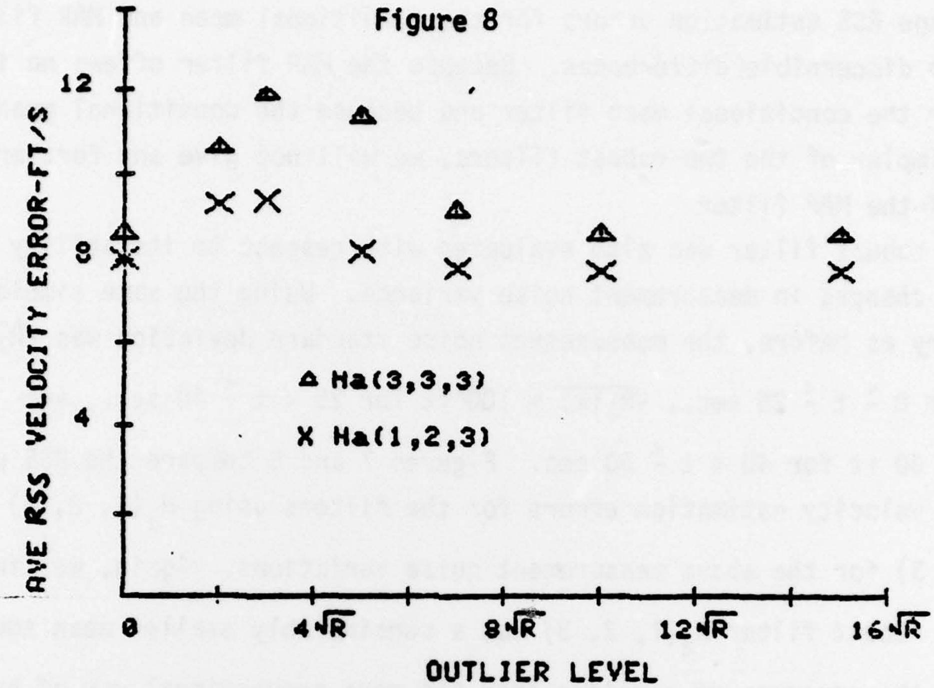


Figure 8



A Monte Carlo evaluation of the robust, conditional mean smoother was performed using the same simulated trajectory as used for the filter evaluation. A smoothing interval of one sec. or 20 samples was used. The Monte Carlo sample size for the smoother evaluation is N = 10. The measurement noise standard deviation was increased so that $\sqrt{R_1(k)}$ = 50 ft. The outlier contamination is the same as for filter evaluation, i.e., a contamination of 8.8% and an average outlier run length of three. Figures 9 and 10 compare the average RSS position and average RSS velocity errors for the two smoothers with $H_a(2, 3, 4)$ and $H_a(4, 4, 4)$ for various levels of outliers. As in the filtering case we see that $H_a(2, 3, 4)$ results in a significant decrease in estimation error from the smoother using the more conventional $H_a(4, 4, 4)$ method of handling outliers.

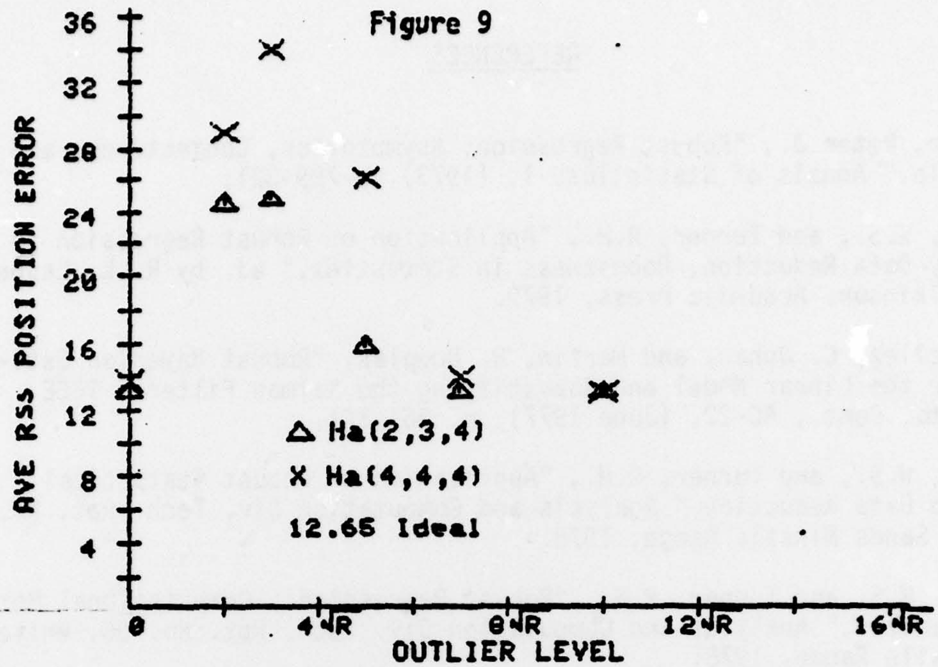


Figure 9

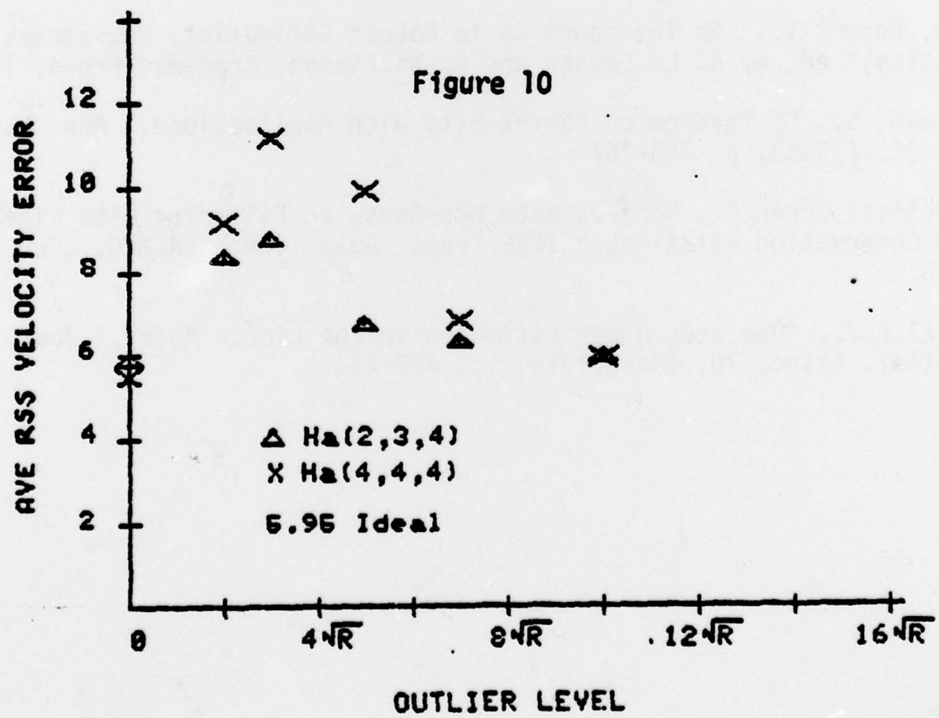


Figure 10

REFERENCES

1. Huber, Peter J., "Robust Regression: Asymptotics, Conjectures, and Monte Carlo," Annals of Statistics, 1, (1973), p-799-821.

2. Agee, W.S., and Turner, R.H., "Application of Robust Regression to Trajectory Data Reduction, Robustness in Statistics," ed. by R. L. Launer and G. Wilkinson, Academic Press, 1979.

3. Masreliez, C. Johan, and Martin, R. Douglas, "Robust Bayesian Estimation for the Linear Model and Robustifying the Kalman Filter," IEEE Trans. Auto. Cont., AC-22, (June 1977), p. 361-371.

4. Agee, W.S., and Turner, R.H., "Application of Robust Statistical Methods to Data Reduction," Analysis and Computation Div. Tech. Rpt. No. 65, White Sands Missile Range, 1978.

5. Agee, W.S. and Turner, R.H., "Robust Regression: Computational Methods for M-Estimates," Analysis and Computation Div. Tech. Rpt. No. 66, White Sands Missile Range, 1978.

6. Hampel, Frank R, "The Influence Curve and its Role in Robust Estimation," J. Amer. Statist. Assoc., 69 , (June 1974), p. 383-393.

7. Hogg, Robert V., "An Introduction to Robust Estimation, Robustness in Statistics," ed. by R. L. Launer and G. Wilkinson, Academic Press, 1979.

8. Sherman, S., "A Theorem on Convex Sets with Applications," Ann. Math Statist., 26, (1955), p. 763-767.

9. Masreliez, Johan C., "Approximate Non-Gaussian Filtering with Linear State and Observation Relations," IEEE Trans. Auto. Cont., AC-20, (Feb 1975), p. 107-110.

10. Bickel,P.J., "One Step Huber Estimates in the Linear Model," Jour. Amer. Statist. Assoc, 70, (June 1975), p. 428-434.